

Zur Beurteilung von Konjunkturprognosen: eine Auswertung von Prognosen zur Entwicklung des BIP, der Industrieproduktion und der privaten Konsumausgaben in zwölf Industrienationen

Spiwoks, Markus; Scheier, Johannes; Hein, Oliver

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

SSG Sozialwissenschaften, USB Köln

Empfohlene Zitierung / Suggested Citation:

Spiwoks, M., Scheier, J., & Hein, O. (2011). *Zur Beurteilung von Konjunkturprognosen: eine Auswertung von Prognosen zur Entwicklung des BIP, der Industrieproduktion und der privaten Konsumausgaben in zwölf Industrienationen*. (sofia-Diskussionsbeiträge zur Institutionenanalyse, 11-1). Darmstadt: Hochschule Darmstadt, FB Gesellschaftswissenschaften und Soziale Arbeit, Sonderforschungsgruppe Institutionenanalyse (sofia). <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-391905>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Zur Beurteilung von Konjunkturprognosen

Eine Auswertung von Prognosen zur Entwicklung des BIP,
der Industrieproduktion und der privaten Konsumausgaben
in zwölf Industrienationen

Markus Spiwoks – Johannes Scheier – Oliver Hein

sofia-Diskussionsbeiträge 11-1, Darmstadt 2011

ISBN: 978-3-941672-03-1

Sofia-Diskussionsbeiträge
zur Institutionenanalyse
Nr. 11-1

ISSN 1437 – 126X
ISBN 978-3-941672-03-1

Inhaltsverzeichnis

1	Einleitung	2
2	Datenbasis	5
3	Methodik	7
4	Ergebnisse	15
5	Fazit	25
	Literatur	27

Tabellenverzeichnis

Tabelle 1: Verwendeter Datenbestand aus der Zeitschrift Consensus Forecasts	6
Tabelle 2: Ergebnisse des Vergleichs mit der naiven Prognose auf der Basis des mittleren absoluten Prognosefehlers bezogen auf die naive Prognose ($MRAPn$) für ausgewählte Prognosehorizonte.	16
Tabelle 3: Ergebnisse des Tests auf Informationszuwachs für ausgewählte Prognosegütemaße und GOVA-Koeffizienten für ausgewählte Prognosehorizonte.	18
Tabelle 4: Ergebnisse des Tests auf Unverzerrtheit in der Mincer-Zarnowitz-Variante	19
Tabelle 5: Gegenüberstellung von Prognoseerfolgen (+) und Prognosemisserfolgen (–) nach ausgewählten Beurteilungsmaßstäben.....	20

1

Einleitung

Konjunkturprognosen sind unverzichtbar. Sowohl das Regierungshandeln als auch viele Unternehmensentscheidungen hängen wesentlich von der erwarteten konjunkturellen Entwicklung ab.

Die Abschätzung der künftigen Steuereinnahmen, die Prognose der notwendigen Sozialausgaben und damit natürlich die gesamte Finanzpolitik eines Staates hängen wesentlich von der erwarteten konjunkturellen Entwicklung ab. Sofern sich eine Regierung dafür verantwortlich fühlt, steuernd in die konjunkturelle Entwicklung einzugreifen, ist sie ebenfalls auf Konjunkturprognosen angewiesen. Auch die Notenbanken erstellen in aller Regel aufwendige Konjunkturprognosen, beispielsweise um eine drohende Überhitzung des Wirtschaftswachstums und somit eine Inflationsgefahr frühzeitig abwenden zu können.

Unternehmen des produzierenden Gewerbes orientieren sich an Konjunkturprognosen, um ihren künftigen Investitions- und Personalbedarf planen zu können oder um sich rechtzeitig den Zugang zu Rohstoffen zu sichern. Für den Bankensektor spielen Prognosen des Wirtschaftswachstums eine besonders wichtige Rolle. Konjunkturprognosen sind notwendig, damit die Banken die Wahrscheinlichkeit für den Ausfall von aktuellen und künftigen Kreditnehmern abschätzen können. Nur so können die Banken einen angemessenen Risikoaufschlag in der Zinsmarge berücksichtigen. Konjunkturprognosen sind auch für das Investmentgeschäft von erheblicher Bedeutung. Die konjunkturelle Entwicklung entfaltet wesentlichen Einfluss auf die Aktien- und Anleihenkurse. Finanzmarktprognosen basieren somit in aller Regel auch auf Konjunkturprognosen.

So wichtig wie Konjunkturprognosen sind, so umstritten sind sie zugleich auch. Das hängt insbesondere damit zusammen, dass der Prognoseerfolg häufig hinter den hochgesteckten Erwartungen zurückbleibt. Die Prognose der konjunkturellen Entwicklung ist eine schwierige Aufgabe. Eine Volkswirtschaft ist ein komplexes Gebilde, das sich darüber hinaus permanent verändert. Evolutionäre Wissenschaftsgegenstände sind besonders schwierig zu ergründen. Das musste die Ökonomen-Gemeinde 2008 im Zuge der Finanzmarktkrise und der darauf folgenden Wirtschaftskrise erneut erkennen.

Die eingeschränkten Möglichkeiten für die Erstellung zuverlässiger Konjunkturprognosen stehen in starkem Kontrast zu den häufig vollkommen überzogenen Ansprüchen der Nutzer dieser Prognosen. Muth (1961) hat mit seiner Hypothese rationaler Erwartungen einen vielleicht entscheidenden Beitrag zu der unrealistischen Einschätzung der Möglichkeiten von Konjunkturprognosen geleistet. Die Hypothese rationaler Erwartungen geht davon aus, dass es bei Berücksichtigung aller relevanten Informationen möglich ist, einen ökonomischen Sachverhalt weitgehend zutreffend zu prognostizieren. Lediglich unvorhersehbare Zufallseinflüsse (das so genannte weiße Rauschen) tragen dazu bei, dass die Prognostiker nicht jede Entwicklung vollkommen zutreffend vorhersagen können.

Ob die Prognosefehler wirklich nur auf unsystematische Zufallseinflüsse zurückzuführen sind, wird in der Regel mit dem Test auf Unverzerrtheit überprüft. Sind die Prognosen nicht unverzerrt, so gelten sie auch nicht als rational.

Die Nutzer von Prognosen fordern von den Experten vernünftige und sinnvolle Prognosen. Es ist nicht immer leicht zu vermitteln, dass vernünftige und sinnvolle Prognosen nicht zwangsläufig auch *rational* im Sinne von Muth (1961) sein müssen. Denn Muth setzt voraus, dass das relevante Modell der Ökonomie bekannt ist, dass also das Funktionieren der Volkswirtschaften weitgehend ergründet ist. Diese Annahme ist aber offensichtlich unzutreffend. Die Bedingungen für wirtschaftliches Wachstum sind nur in Grundzügen bekannt. Viele Einflussfaktoren werden bisher nicht, nicht ausreichend oder nicht korrekt berücksichtigt. In einer Welt unvollkommener Kenntnis über die Zusammenhänge und Wirkmechanismen in der Volkswirtschaft muss man bei der Beurteilung von Prognostikern andere Maßstäbe als den Test auf Unverzerrtheit anlegen.

Die vorliegende Studie verfolgt vier Ziele:

1. Die Studie soll einen kurzen Überblick über die wichtigsten Verfahren der Prognosegütemessung bei makroökonomischen Prognosen geben.
2. Es sollen zwei neue Verfahren zur Beurteilung makroökonomischer Prognosen vorgestellt werden: Der Test auf Informationszuwachs und der GOVA-Koeffizient. Unseres Erachtens handelt es sich um besonders geeignete Ansätze, weil sie realistische Anforderungen an die Prognostiker stellen.
3. Es soll gezeigt werden, dass Prognoseerfolge und Prognosemisserfolge mit dem Test auf Unverzerrtheit nicht angemessen differenziert werden können.

4. Die Studie soll einen Überblick über den Erfolg von Konjunkturprognosen in zwölf Industrienationen geben. Im Einzelnen handelt es sich um Konsensprognosen zur Entwicklung des BIP, der Industrieproduktion und der privaten Konsumausgaben in den USA, Japan, Deutschland, Frankreich, Großbritannien, Italien, Spanien, Kanada, Niederlande, Schweiz, Schweden und Norwegen.

Das nächste Kapitel stellt die verwendeten Prognosedaten näher vor. Die Methodologie steht im Mittelpunkt des übernächsten Kapitels. Im vorletzten Kapitel präsentieren wir die empirischen Ergebnisse. Das letzte Kapitel fasst die wichtigsten Ergebnisse der Studie zusammen.

2 Datenbasis

Im Mittelpunkt der Untersuchung steht die Beurteilung von Konjunkturprognosen. Im Einzelnen handelt es sich um Prognosen zum Wachstum des BIP, der Industrieproduktion und der privaten Konsumausgaben. Die Prognosen beziehen sich jeweils auf die prozentuale Veränderung gegenüber dem letzten Kalenderjahr. Die Prognosen stammen aus der Zeitschrift Consensus Forecasts. Dieses Magazin veröffentlicht seit Oktober 1989 in monatlichem Rhythmus Prognosen zu verschiedenen makroökonomischen Größen. Consensus Forecasts befragt Analysten und Wissenschaftler aus Banken, Versicherungen, Investmentgesellschaften, Forschungsinstituten und Verbänden. Das Magazin veröffentlicht alle Prognosen der Experten und fasst diese schließlich zu Konsensprognosen zusammen.

Die Prognosen beziehen sich auf die prozentuale Veränderung des laufenden Jahres im Vergleich zum Vorjahr und auf die prozentuale Veränderung des folgenden Jahres im Vergleich zum laufenden Jahr. Die Experten geben ihre Einschätzungen jeweils zu Monatsbeginn ab. Die Prognosen beziehen sich jeweils auf das Ende des Jahres, weil erst dann die prozentuale Veränderung gegenüber dem vorherigen Kalenderjahr ermittelt werden kann. Da die Experten jeden Monat diese Prognosen abgeben, resultieren für jeden Prognosegegenstand 24 unterschiedliche Prognosehorizonte. Gibt man beispielsweise Anfang Januar 2011 eine Prognose für das folgende Jahr ab (Stichtag 31.12.2012), so weist diese Prognose einen Prognosehorizont von 24 Monaten auf. Gleichzeitig gibt man eine Prognose für das laufende Jahr ab (Stichtag 31.12.2011). Diese Prognose weist somit einen Prognosehorizont von zwölf Monaten auf. Gibt man Anfang Februar wieder je eine Prognose zum folgenden Jahr und zum laufenden Jahr ab, verkürzen sich die Prognosehorizonte auf 23 Monate beziehungsweise auf elf Monate. Gibt man die zwei Prognosen Anfang Dezember ab, so betragen die Prognosehorizonte noch 13 Monate beziehungsweise einen Monat. Für jeden Prognosegegenstand und jedes prognostizierte Jahr existieren somit 24 Konsensprognosen mit den Prognosehorizonten 24 Monate bis einen Monat. Die Experten beginnen also jeweils zwei Jahre vor Ende eines Kalenderjahres, ihre Prognosen abzugeben. Sie geben jeden Monat neue Prognosen ab, bis nach 24 Monaten das prognostizierte Ereignis eintritt. Diese besondere Struktur der Prognosedaten eröffnet interessante Analysemöglichkeiten.

Consensus Forecasts veröffentlicht bereits seit Oktober 1989 Prognosen für die USA, Japan, Deutschland, Frankreich, Großbritannien, Italien und Kanada. Ab Dezember 1994 finden auch Spanien, Schweden und die Niederlande Berücksichtigung. Im Juni 1998 kommen schließlich noch Norwegen und die

Schweiz hinzu. Wir berücksichtigen in unserer Studie nur die Jahre, für die alle 24 Prognosehorizonte vorliegen. Die Analyse der ersten Ländergruppe umfasst somit die Jahre 1991 bis 2009, die der zweiten Ländergruppe von 1996 bis 2009 und die der dritten Ländergruppe 2000 bis 2009 (siehe Tabelle 1).

Tabelle 1: Verwendeter Datenbestand aus der Zeitschrift Consensus Forecasts

Land	Beginn der Veröffentlichung	Erstes vollständig prognostiziertes Jahr	Letztes vollständig prognostiziertes Jahr	Anzahl der analysierten Prognosen
USA	Okt. 1989	1991	2009	1.368
Japan	Okt. 1989	1991	2009	1.368
Deutschland	Okt. 1989	1991	2009	1.368
Frankreich	Okt. 1989	1991	2009	1.368
Großbrit.	Okt. 1989	1991	2009	1.368
Italien	Okt. 1989	1991	2009	1.368
Spanien	Dez. 1994	1996	2009	1.008
Kanada	Okt. 1989	1991	2009	1.368
Niederlande	Dez. 1994	1996	2009	1.008
Schweiz	Juni 1998	2000	2009	720
Schweden	Dez. 1994	1996	2009	1.008
Norwegen	Juni 1998	2000	2009	720
Gesamt				14.040

Für die erste Ländergruppe liegen also je 24 Prognosen für jedes der 19 betrachteten Jahre (von 1991 bis 2009) vor. Das sind insgesamt 456 Prognosen. Da wir bei jedem Land jeweils drei Prognosegegenstände betrachten (BIP, Industrieproduktion, privater Konsum), führt dies zu 1.368 Prognosen pro Land der ersten Ländergruppe. Insgesamt werten wir 14.040 Konsensprognosen aus.

3 Methodik

Prognosen sollten im günstigsten Fall unverzerrt sein (Diebold, 2007). Das bedeutet, dass der Erwartungswert der Abweichungen zwischen den Prognosen und den tatsächlichen Ereignissen Null sein sollte. Außerdem sollte die Zeitreihe der Residuale keine systematischen Komponenten enthalten, sondern zufällig verteilt sein.

Am häufigsten findet man den Test auf Unverzerrtheit in Form der Mincer-Zarnowitz-Regression (Ball, 1962, Mincer und Zarnowitz, 1969). Seien A_t der tatsächliche Wert, P_t die dazugehörige Prognose und u_t das Residual zum Zeitpunkt t . Die Mincer-Zarnowitz-Regression hat dann folgende Gestalt:

$$(1) \quad A_t = \alpha + \beta P_t + u_t$$

Basierend auf dieser Gleichung gelten die Prognosen als unverzerrt, wenn sich α nicht signifikant von 0 und β sich nicht signifikant von 1 unterscheiden. Zudem darf der Fehlerterm u nicht autokorreliert sein.

Wir stehen dem Test auf Unverzerrtheit allerdings sehr kritisch gegenüber. Er wurzelt in der Theorie rationaler Erwartungen. Damit Prognosen sich als unverzerrt erweisen können, müssen mehrere Voraussetzungen erfüllt sein: 1. Das Geschehen in der realen Volkswirtschaft muss stabile Strukturen aufweisen. 2. Das Model von dieser Volkswirtschaft muss zutreffend und eindeutig sein. 3. Die Prognostiker müssen stets alle relevanten Informationen angemessen berücksichtigen.

In der euphorischen Anfangszeit der neoklassischen Theorie mag man noch bereit gewesen sein, solche Annahmen zumindest als näherungsweise zutreffend anzuerkennen. Nicht zuletzt die Banken- und Finanzmarktkrise 2008, die darauf folgende Wirtschaftskrise 2009, die darauf folgende Staatsschuldenkrise 2010 und die bereits drohende Währungskrise haben viele Wirtschaftswissenschaftler dazu gebracht, an den bisherigen Modellen zu zweifeln.¹ Die Behavioral Economics haben vielfältige Belege dafür gefunden, dass selbst hochbegabte Subjekte nicht in der Lage sind, immer alle relevanten Informationen angemessen auszuwerten. Die Transaktionskostentheorie zeigt, dass dies – selbst wenn es denn möglich wäre – nicht immer sinnvoll ist. Die Evolutionsökonomik hat tiefgreifende Zweifel geweckt, ob die Volkswirtschaften

¹ Nur wenige Ökonomen, wie beispielsweise Thomas Sargent von der New York University oder Robert Barro von der Harvard University, erklärten auf der Jahrestagung der American Economic Association (AEA) 2010 in Atlanta ihr unbeirrtes Festhalten an den Grundlagen der Neoklassik wie zum Beispiel der Theorie rationaler Erwartungen (Storbeck, 2010).

überhaupt dauerhafte, stabile Strukturen aufweisen. Vor diesem Hintergrund erscheint es uns vollkommen unangemessen, von den Prognostikern unverzerrte Prognosen zu verlangen.

Wir haben den Test auf Unverzerrtheit in der Mincer-Zarnowitz-Variante trotzdem in unserer Studie berücksichtigt. Wir wollen aufzuzeigen, dass dieser sehr verbreitete Ansatz zu ganz anderen Ergebnissen führt, als die von uns vorgeschlagenen Verfahren der Prognosebeurteilung.

Der Effizienztest von Holden und Peel (1990) baut unmittelbar auf der Mincer-Zarnowitz-Regression auf.

$$(2) \quad A_t = \alpha + \beta P_t + \gamma X_t + u_t$$

X_t ist eine beliebige zur Verfügung stehende Information. Basierend auf Gleichung (2) gelten die Prognosen als effizient, wenn sich α und γ nicht signifikant von 0 und β nicht signifikant von 1 unterscheiden. Auch hier darf der Fehlerterm nicht autokorreliert sein.

Wenn γ signifikant von Null unterschiedlich ist, dann wurde die Information X nicht angemessen in den Prognosen berücksichtigt. Bei diesem Ansatz greifen im Wesentlichen dieselben Kritikpunkte wie beim Test auf Unverzerrtheit.

Eine andere übliche Variante des Effizienztests betrachtet, ob die letzten tatsächlichen Werte vor Abgabe einer Prognose systematischen Einfluss auf den Prognosefehler haben (Simon, 1989). Ist dies der Fall, so stecken in den tatsächlichen Ereignissen Informationsbestandteile, die in den Prognosen unberücksichtigt bleiben. Seien A_t wieder das aktuelle Ereignis, P_t die entsprechende Prognose und u_t das Residual zum Zeitpunkt t . Des Weiteren erfolgt eine Berücksichtigung des Prognosehorizonts h .

$$(3) \quad A_t - P_t = \alpha + \sum_{i=1}^h \beta_i A_{t-h-i} + u_t$$

Wenn die zur Verfügung stehenden Informationen effizient ausgewertet wurden, dürfen die Prognosefehler nicht signifikant mit den LAG-Variablen der tatsächlichen Ereignisse korreliert sein. Ob dies der Fall ist, wird mit Hilfe eines F-Tests ermittelt.

Diesen Ansatz beurteilen wir positiver. Die tatsächlichen Ereignisse der jüngsten Vergangenheit vor einer Prognoseabgabe müssen von einem Prognostiker zur Kenntnis genommen werden. Wenn sich ein systematischer Zusammenhang zwischen den zurückliegenden und den künftigen Ereignissen herstellen lässt, dann sollte man von einem Prognostiker erwarten, dass er diese Möglichkeit nutzt.

Der Vorzeichentest (Merton, 1981; Henriksson und Merton, 1981) ist ein weiteres weit verbreitetes Instrument zur Beurteilung von Prognosen. Bei diesem bekannten Verfahren wird nicht das Ausmaß einer prognostizierten Veränderung thematisiert. Es wird lediglich überprüft, ob die Richtung der Prognosen stimmt. Die Prognosen werden dann in eine 2×2 -Matrix eingeteilt. Dabei unterscheidet man einerseits, ob eine Wachstumsbeschleunigung oder eine Wachstumsverlangsamung vorhergesagt wurde. Andererseits unterscheidet man, ob eine Wachstumsbeschleunigung oder eine Wachstumsverlangsamung eingetreten ist. Die Hauptdiagonale in der 2×2 -Matrix kennzeichnet die in der Entwicklungstendenz zutreffenden Prognosen. Die Nebendiagonale kennzeichnet die in der Entwicklungstendenz unzutreffenden Prognosen. Mittels eines Chi-Quadrat-Tests wird nun überprüft, ob sich die Häufigkeitsverteilung in den 4 Feldern signifikant von einer Zufallsprognose unterscheidet (vgl. Diebold und Lopez, 1996; Joutz und Stekler, 2000). Sollte dies der Fall sein, so muss weiterhin geprüft werden, ob die Prognosen besser oder schlechter als die Zufallsprognose abschneiden.

Wir halten diesen Ansatz für adäquat. Er überprüft, ob die Prognostiker unwissend sind und lediglich raten oder ob sie zumindest einen groben Einblick in den prognostizierten Sachverhalt gewonnen haben. Hier werden also keine überzogenen Ansprüche an die Experten formuliert.

Der „Goldstandard“ der Prognosegütemessung ist nach wie vor der Vergleich zur naiven Prognose. Angenommen eine Blackbox produziert in regelmäßigen Abständen ein quantitatives Ereignis. Soll nun, in völliger Unkenntnis der in der Blackbox stattfindenden Prozesse, eine Prognose über den weiteren Verlauf der Zeitreihe abgegeben werden, so spricht eben so viel für einen steigenden wie für einen fallenden Verlauf in der Zukunft. Es erscheint also sinnvoll, von einer unveränderten Situation in der Zukunft auszugehen (naive Prognose). Der französische Mathematiker Pierre Simon Laplace (1814) hat diese Überlegung als „Prinzip des unzureichenden Grundes“ in die Literatur eingeführt. Seitdem gilt die naive Prognose als absoluter Nullpunkt der Prognosequalität, da diese mühelos und ohne jede Kenntnis der prognostizierten Zusammenhänge erreichbar ist. Schon ein rudimentäres Verständnis der zu Grunde liegenden Prozesse müsste zu einer deutlich besseren Prognosegüte als bei der naiven Prognose führen.²

Es gibt vielfältige Möglichkeiten, um einen Vergleich zur naiven Prognose vorzunehmen. Zunächst lassen sich einfache Prognosegütemaße jeweils für die Prognosen der Experten und für die naiven Prognosen berechnen. Die gebräuchlichsten einfachen Prognosegütemaße sind der mittlere Prognosefehler

² Vgl. Scheier, J. / Spiwoks, M. (2006), S. 63.

(MP), Root Mean Squared Error (RMSE), der mittlere absolute Prognosefehler (MAP), Median Absolute Error (MdAE), Mean Absolute Percentage Error (MAPE), Median Absolute Percentage Error (MdAPE), Symmetric Mean Absolute Percentage Error (sMAPE), Symmetric Median Absolute Percentage Error (sMdAPE), Mean Relative Absolute Error (MRAE), Median Relative Absolute Error (MdRAE) und Geometric Mean Relative Absolute Error (GMRAE) (De Gooijer und Hyndman, 2006; Mathews und Diamantopoulos, 1994).

Etwas bequemer ist die Handhabung des Vergleichs zur naiven Prognose, wenn er bereits implizit in dem Prognosegütemaß vorgesehen ist. Theil's U ist ein bekanntes Beispiel dafür (Theil, 1971). Will man auf eine überproportionale Gewichtung von großen Prognosefehlern verzichten, dann bietet sich der mittlere absolute Prognosefehler bezogen auf die naive Prognose (*MRAPnP*) an.

$$(4) \quad MRAPnP = \frac{\sum_{t=1}^T |P_t - A_t|}{\sum_{t=1}^T |A_{t-h} - A_t|}$$

Dieser besteht im Zähler aus der Summe der absoluten Prognosefehler (P_t = Prognose zum Zeitpunkt t , A_t = tatsächliches Ereignis zum Zeitpunkt t). Diese wird durch die Summe der Prognosefehler der naiven Prognose geteilt, wobei A_{t-h} den tatsächlichen Wert zum Zeitpunkt $t-h$ darstellt (h =Prognosehorizont).

Fair und Shiller (1990) haben gezeigt, dass diese Ansätze in bestimmten Konstellationen zu fragwürdigen Ergebnissen führen können. Deshalb findet der Diebold Mariano Test for Forecast Encompassing (Diebold und Mariano, 1995) immer häufiger Verwendung. Ausgangspunkt ist hierbei die Vorhersage einer Situation P_m durch zwei konkurrierende Prognosemodelle i und j :

$$(5) \quad P_m = (1 - \lambda) P_{i,m} + \lambda P_{j,m}$$

wobei $0 \leq \lambda \leq 1$. Für $\lambda = 0$ spricht man von einer Umfassung der vom Modell j generierten Prognosen durch die Vorhersagen des Modells i . Das Modell j kann in diesem Fall nichts zu einer optimalen Kombinationsprognose aus beiden Modellen beitragen, da es über die in Modell i bereits enthaltenen Informationen hinaus keine weiteren nützlichen Zusatzinformationen beisteuert. Harvey, Leybourne und Newbold (1998) stellen eine Möglichkeit vor, die Nullhypothese $H_0: \lambda = 0$ gegen die Alternative $H_1: \lambda > 0$ zu testen. Bei einer Ablehnung der Nullhypothese enthalten beide Modelle unterschiedliche Informationen mit Zukunftsgehalt, so dass sie zur Bildung der optimalen

Kombinationsprognose P_m herangezogen werden können. Betrachtet man nun die naiven Prognosen als Modell i und die Expertenprognosen als Modell j , so kann ermittelt werden, ob der Informationsgehalt der Expertenprognosen signifikant über den Informationsgehalt der naiven Prognosen hinausgeht.

Wir halten den Vergleich zur naiven Prognose für sinnvoll. Prognoseverfahren, die nicht zu besseren Ergebnissen führen als die entsprechenden naiven Prognosen, sind zweifelsfrei untauglich. Der Diebold Mariano Test for Forecast Encompassing bietet eine besonders elegante Möglichkeit, um einen Vergleich zur naiven Prognose vorzunehmen. Allerdings haben wir Zweifel, ob es vernünftig ist, immer und überall nur statistisch signifikant bessere Ergebnisse als Erfolg anzuerkennen. Beispielsweise fragt bei der Olympiade niemand danach, ob der Goldmedaillengewinner beim 100-Meter-Lauf statistisch signifikant schneller laufen kann als der Zweitplatzierte.³ Wenn es darum geht, keine übertriebenen Anforderungen an die Prognostiker zu stellen, so kommen Vergleiche auf der Basis einfacher Prognosegütemaße nach wie vor in Betracht. Wir verwenden den Mittleren absoluten Prognosefehler bezogen auf die naive Prognose ($MRAPnP$), weil es sich dabei um ein einfaches, übersichtliches sowie ein leicht und eindeutig zu interpretierendes Prognosegütemaß handelt.

Allerdings bietet es sich nicht an, dieses Prognosegütemaß für alle 24 Prognosehorizonte zu ermitteln. Die tatsächlichen Ereignisse treten nicht monatlich, sondern nur jahresweise auf. Dementsprechend gibt es pro Jahr auch nur eine naive Prognose, die als Vergleichsmaßstab beim $MRAPnP$ verwendet werden kann. Die Prognosen, die im Januar abgegeben werden, können sinnvoll mit den naiven Prognosen (tatsächliches Ereignis des Vorjahres) verglichen werden. Deshalb werten wir die Prognosen mit 24 und mit 12 Monaten Prognosehorizont aus. Hier könnte allerdings eingewendet werden, dass Anfang Januar die Daten zum tatsächlichen Wirtschaftswachstum des Vorjahres noch nicht vorliegen. Üblicherweise dauert es bis März, bis die ersten weitgehend verlässlichen Zahlen vorliegen. Deshalb werten wir außerdem die Prognosen, die Anfang April abgegeben werden, mit dem $MRAPnP$ aus. Das betrifft somit die Prognosen mit 21 und mit neun Monaten Prognosehorizont. Da wir überzogene Ansprüche an die Prognostiker vermeiden wollen, erscheint die Betrachtung der Prognosen von Anfang April sogar sinnvoller. Denn die kostengünstige Alternative, sich stets auf die naiven Prognosen zu stützen, steht schließlich erst ab Anfang April zur Verfügung.

³ Es wäre ja ohne Weiteres möglich, auch bei der Beurteilung sportlicher Leistungen statistische Signifikanztests zu etablieren: Die Sprinter müssten 30 Mal gegeneinander antreten. Aus den durchschnittlichen Zeiten und der Streuung der Einzelergebnisse ließe sich dann ermitteln, ob der Erste signifikant schneller laufen kann als der Zweitplatzierte. Eine solche Vorgehensweise dürfte den meisten Sportbegeisterten jedoch als weltfremd und unsinnig erscheinen.

Je weiter (näher) nicht-deterministische Ereignisse in der Zukunft liegen, desto schwieriger (leichter) sind sie in aller Regel zu prognostizieren. Das ist darauf zurückzuführen, dass im Zeitverlauf Ereignisse zu beobachten sind, die zu einer besseren Beurteilung des Sachverhaltes beitragen können (Poulizac, Weale und Young, 1996, Ackert und Hunter, 1994, S. 390-391, Döpke und Fritsche, 2006, S. 789). Den Ausgang eines Fußballspiels kann man in der 80. Spielminute weitaus besser prognostizieren als bei Anpfiff des Spiels. Hat eine Fußballmannschaft bis zur 80. Spielminute zwei Spieler durch rote Karten verloren und liegt zugleich mit 0:6 Toren zurück, dann ist es nicht mehr schwierig, diese Mannschaft als Verlierer zu prognostizieren. Müsste man den Ausgang von 100 Fußballspielen jeweils zu Spielbeginn prognostizieren, so würde man gewiss viel häufiger falsch liegen, als wenn man den Ausgang von 100 Fußballspielen jeweils in der 80. Minute prognostizieren würde. Voraussetzung dafür ist nur, dass man über einige rudimentäre Grundkenntnisse des Fußballs verfügt und die relevanten Ereignisse (Tore, Platzverweise, etc.) bis zur 80. Spielminute zur Kenntnis nimmt.

Von Wirtschaftsexperten sollte man das gleiche verlangen können. Sie sollten zumindest über rudimentäre Grundkenntnisse wirtschaftlicher Prozesse verfügen⁴ und zumindest die wichtigsten der relevanten Ereignisse zur Kenntnis nehmen. Sind diese beiden Voraussetzungen erfüllt, dann sollten die Prognosen im langjährigen Durchschnitt umso genauer sein, je kürzer der Prognosehorizont ist.

Man könnte vermuten, dass Wirtschaftsprognostiker diese Anforderung üblicherweise ohne jede Anstrengung und ohne jede Schwierigkeit bewältigen. Tatsächlich zeigt sich aber, dass beispielsweise Finanzmarktanalysten diesem Anforderungsprofil keineswegs immer entsprechen. So zeigt Spiwoks (2009), dass der Prognosefehler bei der Vorhersage der Gewinnentwicklung US-amerikanischer Unternehmen mit abnehmendem Prognosehorizont nicht zurückgeht. Die Prognosen mit einem Prognosehorizont von nur einem Monat sind im Durchschnitt schlechter als die Prognosen mit einem Prognosehorizont von 24 Monaten. Dies ist besonders überraschend, weil die Unternehmen quartalsweise ihre Gewinne veröffentlichen, so dass bei kurzen Prognosehorizonten bereits wesentliche Bestandteile des zu prognostizierenden Ereignisses bekannt sind. Offensichtlich sind die Analysten aber nicht bereit, diese Zwischenergebnisse in ihren Prognosen zu berücksichtigen.

Wir schlagen daher einen Test auf Informationszuwachs vor. Dieser neue und zugleich sehr einfache Ansatz überprüft, ob die Prognostiker bereit und in der Lage sind, relevante Ereignisse während eines sich allmählich verkürzenden

⁴ Mehr ist bei dem aktuellen Forschungsstand der Wirtschaftswissenschaften vielleicht auch gar nicht möglich.

Prognosehorizontes zur Kenntnis zu nehmen. Es wird also überprüft, ob der Prognosefehler mit abnehmendem Prognosehorizont statistisch signifikant zurückgeht.

Die Prognosefehler werden als absoluter Prognosefehler (AP) oder als absoluter prozentualer Prognosefehler (APP) für jeden Prognosehorizont berechnet. Anschließend wird der Korrelationskoeffizient zwischen Prognosehorizont und Prognosefehler geschätzt. Weist der Korrelationskoeffizient statistisch signifikant ein positives Vorzeichen auf, so verringert sich der Prognosefehler erkennbar mit abnehmendem Prognosehorizont. Eine wichtige Mindestanforderung an die Prognostiker ist in diesem Fall als erfüllt anzusehen.

Der Test auf Informationszuwachs kann der Gruppe der Effizienztests zugeordnet werden. Denn es wird überprüft, ob relevante Informationen angemessene Berücksichtigung in den Prognosen finden.

Wenn Prognosen stärker mit den tatsächlichen Werten zum Zeitpunkt ihrer Erstellung als mit den Entwicklungen zu ihren Gültigkeitszeitpunkten korreliert sind, spricht man von gegenwartsorientierter Verlaufsanpassung (GOVA). Hierbei erfolgt eine fortlaufende Anpassung der Prognosen an das aktuelle Marktgeschehen, so dass ihr zukunftsorientierter Charakter unter Umständen vollständig verloren geht. Mit Hilfe des GOVA-Koeffizienten lässt sich eine etwaige gegenwartsorientierte Verlaufsanpassung feststellen.

Hierzu wird zunächst das Bestimmtheitsmaß (R^2_A) zwischen den Prognosen und den tatsächlich eingetretenen Werten berechnet (siehe Andres und Spiwoks, 1999, sowie Bofinger und Schmidt, 2003). Anschließend erfolgt die Ermittlung des Bestimmtheitsmaßes zwischen den Prognosen und den Werten des Prognosegegenstands zum Zeitpunkt der Prognoseerstellung (R^2_B). Die Division der beiden Bestimmtheitsmaße ergibt nun den GOVA-Koeffizienten.

$$(6) \quad \text{GOVA-Koeffizient} = \frac{R^2_{\text{Prognosen (Geltungszeitpunkte); tatsächliche Ereignisse}}}{R^2_{\text{Prognosen (Entstehungszeitpunkte); tatsächliche Ereignisse}}} = \frac{R^2_A}{R^2_B}$$

Bei einem GOVA-Koeffizienten < 1 liegt gegenwartsorientierte Verlaufsanpassung vor. In diesem Fall haben die Prognosen den Prognosegegenstand zum Zeitpunkt ihrer Erstellung besser erfasst als zum Zeitpunkt ihrer eigentlichen Gültigkeit. Sie spiegeln die Vergangenheit stärker als die Zukunft wider.

Bei Finanzmarktprognosen treten gegenwartsorientierte Verlaufsanpassungen regelmäßig auf. Spiwoks, Gubaydullina und Hein (2011) untersuchen beispielsweise rund 1.200 Zinsprognosezeitreihen mit rund 160.000 einzelnen Prognosen. Es zeigt sich, dass 98.5% aller Prognosezeitreihen von einer ge-

genwartsorientierten Verlaufsanpassung geprägt sind. Finanzmarktanalysten können oder wollen sich offenbar nicht von der jeweils aktuell vorherrschenden Marktsituation lösen. Sie „prognostizieren“ also vor allem die Gegenwart, aber kaum die Zukunft. Schuh (2001, S. 42), Mehra (2002, S. 21-22), Mankiw, Reis und Wolfers (2003, S. 212), Bowles et al. (2007, S. 18), Andolfatto, Hendry und Moran (2008, S. 407) und Dovern, Fritsche und Slacalek (2009, S. 38-43) geben erste Hinweise darauf, dass gegenwartsorientierte Verlaufsanpassungen durchaus auch bei makroökonomischen Prognosen auftreten können.

Der GOVA-Koeffizient wurde jedoch bisher noch nie zur Analyse makroökonomischer Prognosen angewendet. Das ist sicherlich darauf zurückzuführen, dass der GOVA-Koeffizient nur für die Analyse von Prognosen mit konstantem Prognosehorizont geeignet ist. Wählt man einzelne Prognosehorizonte aus den vorhandenen 24 Prognosehorizonten aus, so lässt sich der GOVA-Koeffizient jedoch auch auf makroökonomische Prognosezeitreihen anwenden. Wir wenden ihn auf die Prognosen mit Prognosehorizonten von zwölf und neun Monaten an. Diese beiden Prognosehorizonte erscheinen besonders geeignet. Im Januar dürfte das abgelaufene Jahr bei den Prognostikern besonders präsent sein, so dass sich möglicherweise eine Neigung zu gegenwartsorientierten Verlaufsanpassungen einstellt. Anfang April sind in aller Regel bereits recht verlässliche vorläufige Daten zum Wirtschaftswachstum des Vorjahres bekannt. Vielleicht verführen gerade diese Daten zu gegenwartsorientierten Verlaufsanpassungen.

Wir analysieren die Konjunkturprognosen also mit vier der vorgestellten Verfahren. Den mittleren absoluten Prognosefehler bezogen auf die naive Prognose ($MRAPnP$), den Test auf Informationszuwachs und den GOVA-Koeffizienten wenden wir an, weil wir diese Instrumente als besonders geeignet betrachten. Den Test auf Unverzerrtheit wenden wir an, um zu zeigen, zu welcher gravierenden Fehleinschätzungen ein ungeeigneter Ansatz zur Beurteilung von Prognosen führen kann.

4 Ergebnisse

Der Vergleich zur naiven Prognose auf der Basis des mittleren absoluten Prognosefehlers bezogen auf die naive Prognose (*MRAPnP*) zeigt, dass die Konjunkturprognostiker bis auf wenige Ausnahmen erfolgreich sind (Tabelle 2). Bei den Prognosehorizonten 24 und 12 Monate sind die Prognosen jeweils zu 33 der 36 Prognosegegenstände (91.7%) genauer als die Alternative der naiven Prognosen.⁵ Beim Prognosehorizont von 21 Monaten weisen die Prognosen der Experten in 34 von 36 Fällen (94.4%) geringere Prognosefehler auf als die entsprechenden naiven Prognosen. Beim kürzesten der betrachteten Prognosehorizonte (neun Monate) sind die Prognosen für alle 36 Prognosegegenstände der naiven Prognose überlegen.

Die naive Prognose stellt den absoluten Nullpunkt der Prognosegüte dar (Fildes und Stekler, 2002, S. 439). Ein Prognostiker muss Prognosen abgeben, die erkennbar besser als die entsprechenden naiven Prognosen sind, um seine Tätigkeit überhaupt rechtfertigen zu können. Denn die naive Prognose steht als kostenlose Alternative zur Verfügung. Insofern scheint es keinen Grund zur Euphorie zu geben, wenn die Konjunkturprognostiker in 136 von 144 Fällen (94.4%) besser als die entsprechenden naiven Prognosen abschneiden.

Um den Erfolg der Konjunkturexperten angemessen würdigen zu können, sollte man sich die Erfolgsquoten von Finanzmarktprognostikern in Erinnerung rufen. In vielen Studien zeigt sich, dass zwischen 80% und 100% der Zeitreihen von Finanzmarktprognosen schlechter als die entsprechenden naiven Prognosen abschneiden (vgl. zu Aktienprognosen: Lakonishok, 1980, Fraser und MacDonald, 1993, Spiwoks, 2004; vgl. zu Zinsprognosen: Brooks und Gray, 2004, Mose, 2005, Spiwoks, Bedke und Hein, 2008, Spiwoks, Bedke und Hein, 2009, Spiwoks, Bedke und Hein, 2010; vgl. zu Wechselkursprognosen: Manzur, 1988, Chinn und Frankel, 1994, Bofinger und Schmidt, 2003). Insofern erscheinen die betrachteten Konjunkturprognosen durchaus als sehr erfolgreich.

⁵ Osterloh (2008) verwendet Theil's U als Beurteilungsmaßstab und kommt für die deutschen BIP-Prognosen mit 24 Monaten Prognosehorizont zu einem ungünstigeren Befund. Die unterschiedlichen Ergebnisse können auf die Auswahl des Beurteilungsmaßstabes zurückgehen. Sie können aber auch in dem unterschiedlichen Betrachtungszeitraum begründet sein. Osterloh wertet Prognosedaten von 1995 bis 2005 aus.

Tabelle 2: Ergebnisse des Vergleichs mit der naiven Prognose auf der Basis des mittleren absoluten Prognosefehlers bezogen auf die naive Prognose (*MRAPnP*) für ausgewählte Prognosehorizonte.

Land	Prognose- gegenstand	<i>MRAPnP</i> 24 Mon.	<i>MRAPnP</i> 21 Mon.	<i>MRAPnP</i> 12 Mon.	<i>MRAPnP</i> 9 Mon.
USA	BIP	0.972	0.941	0.793	0.639
	Industrieproduktion	0.989	0.966	0.808	0.645
	privater Konsum	0.914	0.827	1.012	0.741
Japan	BIP	0.825	0.780	0.736	0.551
	Industrieproduktion	0.741	0.726	0.643	0.432
	privater Konsum	0.932	0.836	0.878	0.633
Deutschland	BIP	0.813	0.772	0.670	0.519
	Industrieproduktion	0.750	0.739	0.620	0.415
	privater Konsum	0.810	0.871	0.952	0.885
Frankreich	BIP	0.785	0.779	0.751	0.502
	Industrieproduktion	0.879	0.859	0.825	0.623
	privater Konsum	0.836	0.860	1.074	0.914
Großbritannien	BIP	0.798	0.769	0.809	0.640
	Industrieproduktion	0.953	0.927	0.708	0.485
	privater Konsum	0.879	0.820	0.800	0.595
Italien	BIP	0.891	0.879	0.728	0.561
	Industrieproduktion	0.951	0.968	0.867	0.612
	privater Konsum	0.943	0.892	0.888	0.746
Spanien	BIP	0.914	0.838	0.931	0.822
	Industrieproduktion	0.846	0.831	0.852	0.555
	privater Konsum	0.744	0.726	0.983	0.813
Kanada	BIP	0.789	0.794	0.774	0.535
	Industrieproduktion	1.087	1.084	1.217	0.877
	privater Konsum	0.737	0.749	0.800	0.669
Niederlande	BIP	0.875	0.853	0.883	0.730
	Industrieproduktion	0.788	0.797	0.767	0.568
	privater Konsum	1.010	0.978	0.970	0.932
Schweiz	BIP	0.773	0.773	0.713	0.653
	Industrieproduktion	0.647	0.653	0.625	0.561
	privater Konsum	0.702	0.685	0.815	0.884
Schweden	BIP	0.823	0.827	0.855	0.673
	Industrieproduktion	0.754	0.781	0.690	0.511
	privater Konsum	0.854	0.871	0.859	0.735
Norwegen	BIP	0.836	0.842	0.756	0.714
	Industrieproduktion	1.218	1.182	0.497	0.525
	privater Konsum	0.809	0.767	0.742	0.689

MRAPnP = mittlerer absoluter Prognosefehler bezogen auf die naive Prognose.

Der Test auf Informationszuwachs vermittelt ebenfalls ein recht positives Bild (Tabelle 3). Betrachtet man die Korrelationskoeffizienten zwischen den Prognosehorizonten und den absoluten Prognosefehlern (AP), so zeigt sich bei 35 von 36 Prognosegegenständen (97.2%) ein positives Vorzeichen. Bei 33 dieser 35 Fälle (94.3%) erweist sich dieses Ergebnis bei einer Irrtumswahrscheinlichkeit von 1% als statistisch signifikant. In den übrigen beiden Fällen zeigt sich eine statistische Signifikanz bei einer Irrtumswahrscheinlichkeit von 5%. Lediglich bei den Prognosen zur Industrieproduktion in Norwegen zeigt sich ein negatives Vorzeichen. Allerdings ist dieses Ergebnis statistisch nicht signifikant.

Tabelle 3: Ergebnisse des Tests auf Informationszuwachs für ausgewählte Prognosegütemaße und GOVA-Koeffizienten für ausgewählte Prognosehorizonte.

Land	Prognose- gegenstand	Korr. PH - <i>AP</i>	Korr. PH - <i>APP</i>	GOVA 12 Mon.	GOVA 9 Mon.
USA	BIP	0.408***	0.215***	1.562	2.102
	Industrieproduktion	0.342***	0.233***	1.606	2.072
	privater Konsum	0.375***	0.177***	1.029	1.432
Japan	BIP	0.312***	0.182***	0.627	1.585
	Industrieproduktion	0.368***	0.272***	1.168	4.904
	privater Konsum	0.233***	0.201***	0.507	1.011
Deutschland	BIP	0.386***	0.127***	3.032	6.248
	Industrieproduktion	0.371***	0.083***	3.099	6.329
	privater Konsum	0.187***	0.227***	1.439	1.957
Frankreich	BIP	0.415***	0.122***	1.178	2.225
	Industrieproduktion	0.335***	0.276***	1.643	2.617
	privater Konsum	0.283***	0.164***	0.760	1.592
Großbritannien	BIP	0.212***	0.089**	1.704	1.860
	Industrieproduktion	0.390***	0.234***	2.680	2.765
	privater Konsum	0.287***	0.205***	1.624	1.900
Italien	BIP	0.417***	0.155***	1.413	1.692
	Industrieproduktion	0.370***	0.251***	1.598	2.969
	privater Konsum	0.280***	0.223***	0.836	1.285
Spanien	BIP	0.281***	0.208***	1.309	1.405
	Industrieproduktion	0.352***	0.257***	0.775	1.104
	privater Konsum	0.250***	0.144**	1.166	1.273
Kanada	BIP	0.425***	0.314***	2.394	2.888
	Industrieproduktion	0.334***	0.156***	3.004	3.403
	privater Konsum	0.283***	0.147***	1.347	1.882
Niederlande	BIP	0.330***	0.171***	1.352	2.033
	Industrieproduktion	0.285***	0.154***	2.749	3.323
	privater Konsum	0.420***	0.218***	1.409	1.295
Schweiz	BIP	0.500***	0.267***	4.501	9.678
	Industrieproduktion	0.366***	0.163**	2.476	2.858
	privater Konsum	0.160**	0.054	0.816	0.643
Schweden	BIP	0.310***	0.133**	0.840	1.229
	Industrieproduktion	0.242***	0.163***	1.573	2.634
	privater Konsum	0.325***	0.077	0.268	0.909
Norwegen	BIP	0.141**	0.102*	1.004	2.524
	Industrieproduktion	-0.043	0.111	1.739	3.372
	privater Konsum	0.391***	0.192**	1.138	1.438

Korr. = Korrelationskoeffizient; PH = Prognosehorizont; *AP* = absoluter Prognosefehler; *APP* = absoluter prozentualer Prognosefehler; GOVA = GOVA-Koeffizient; Signifikanzniveau: 1%***, 5%** und 10%*.

Tabelle 4: Ergebnisse des Tests auf Unverzerrtheit in der Mincer-Zarnowitz-Variante

Land	Prognose- gegenstand	α	β	F-Test p-Wert	DW p-Wert
USA	BIP	0.264	0.890	0.001 ¹	0.000
	Industrieproduktion	-0.402	0.937	0.000	0.000
	privater Konsum	0.272	1.029	0.000 ¹	0.000
Japan	BIP	-0.059	0.650	0.000	0.000
	Industrieproduktion	-1.627	0.803	0.000	0.000
	privater Konsum	0.347	0.484	0.000 ¹	0.000
Deutschland	BIP	-0.064	0.879	0.000	0.000
	Industrieproduktion	-1.101	0.973	0.000	0.000
	privater Konsum	0.453	0.717	0.000 ¹	0.000
Frankreich	BIP	-0.230	0.915	0.000	0.000
	Industrieproduktion	-1.759	0.977	0.000	0.000
	privater Konsum	0.010	0.914	0.001 ¹	0.000
Großbritannien	BIP	-0.259	1.143	0.000 ¹	0.000
	Industrieproduktion	-1.151	0.907	0.000	0.000
	privater Konsum	-0.382	1.249	0.000 ¹	0.000
Italien	BIP	-0.671	1.012	0.000	0.000
	Industrieproduktion	-1.765	1.047	0.000	0.000
	privater Konsum	-0.197	0.832	0.000	0.000
Spanien	BIP	-0.475	1.300	0.000 ¹	0.000
	Industrieproduktion	-2.044	1.028	0.000 ¹	0.000
	privater Konsum	-0.984	1.410	0.000 ¹	0.000
Kanada	BIP	-0.699	1.172	0.000	0.000
	Industrieproduktion	-0.648	0.654	0.000 ¹	0.000
	privater Konsum	0.143	1.010	0.002 ¹	0.000
Niederlande	BIP	0.089	1.045	0.081	0.000
	Industrieproduktion	-0.443	0.940	0.000	0.000
	privater Konsum	0.178	0.984	0.179 ¹	0.000
Schweiz	BIP	0.133	1.014	0.191	0.000
	Industrieproduktion	-0.816	1.065	0.062	0.000
	privater Konsum	0.693	0.581	0.000	0.000
Schweden	BIP	-0.387	1.153	0.125	0.000
	Industrieproduktion	-3.188	1.228	0.000 ¹	0.000
	privater Konsum	0.801	0.688	0.000 ¹	0.000
Norwegen	BIP	0.702	0.419	0.000 ¹	0.000
	Industrieproduktion	-1.535	0.396	0.000	0.000
	privater Konsum	0.465	1.005	0.000	0.000

¹ = Heteroskedastizität kann nicht ausgeschlossen werden, so dass der p-Wert mit robusten Standardfehlern ermittelt wurde; die p-Werte, die sich durch Schätzung mit robusten Standardfehlern geändert haben, sind fett hervorgehoben; DW = Durbin-Watson-Test.

Tabelle 5: Gegenüberstellung von Prognoseerfolgen (+) und Prognoseemisserfolgen (–) nach ausgewählten Beurteilungsmaßstäben.

Land	Prognose- gegenstand	<i>MRAPhP</i> 9 Mon.	Korrel. PH- <i>AP</i>	GOVA 9 Mon.	Test auf Unverzerrth.
USA	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	+	–
Japan	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	+	–
Deutschland	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	+	–
Frankreich	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	+	–
Großbritannien	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	+	–
Italien	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	+	–
Spanien	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	+	–
Kanada	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	+	–
Niederlande	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	+	–
Schweiz	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	–	–
Schweden	BIP	+	+	+	–
	Industrieproduktion	+	+	+	–
	privater Konsum	+	+	–	–
Norwegen	BIP	+	+	+	–
	Industrieproduktion	+	–	+	–
	privater Konsum	+	+	+	–

+ = Prognoseerfolg; – = Prognoseemisserfolg; PH = Prognosehorizont; *AP* = absoluter Prognosefehler; *MRAPhP* = mittlerer absoluter Prognosefehler bezogen auf die naive Prognose; GOVA = GOVA-Koeffizient.

Etwas weniger zwingend stellt sich dieser Zusammenhang dar, wenn man die Korrelationskoeffizienten zwischen den Prognosehorizonten und den absoluten prozentualen Prognosefehlern (*APP*) betrachtet (Tabelle 3). Zwar weisen alle 36 Korrelationskoeffizienten ein positives Vorzeichen auf, aber nur 27 davon (75.0%) sind bei einer Irrtumswahrscheinlichkeit von 1% als statistisch signifikant zu betrachten. Bei einer Irrtumswahrscheinlichkeit von 5% erweisen sich immerhin 32 der 36 positiven Korrelationskoeffizienten (88.9%) als statistisch signifikant. Bei einer Irrtumswahrscheinlichkeit von 10% sind es 33 der 36 Korrelationskoeffizienten (91.7%).

Der Test auf Informationszuwachs zeigt somit nur in 4 von 72 Fällen (5.6%) keinen statistisch signifikanten Zusammenhang. In 68 von 72 Fällen (94.4%) zeigt sich hingegen, dass der Prognosefehler umso geringer ausfällt, je kürzer der Prognosehorizont ist. Die überwiegende Mehrheit der Konjunkturprognostiker ist also offenbar bereit und in der Lage, im Zeitverlauf relevante Informationen aufzunehmen und sinnvoll in ihre Prognosen einzubeziehen.

Es mag auf den ersten Blick wenig beeindruckend wirken, dass Konjunkturexperten ihre Prognosen nicht abgeben, ohne im Laufe der Zeit relevante Fakten zur Kenntnis zu nehmen. Doch zeigt sich beispielsweise bei Finanzmarktprognostikern, dass dies keineswegs als selbstverständlich zu betrachten ist.⁶

Der GOVA-Koeffizient vermittelt ebenso wie der Vergleich zur naiven Prognose und der Test auf Informationszuwachs ein verhältnismäßig positives Bild von den Leistungen der Konjunkturexperten (Tabelle 3). Bei 28 der 36 Prognosezeitreihen mit 12 Monaten Prognosehorizont (77.8%) und bei 34 von 36 Prognosezeitreihen mit 9 Monaten Prognosehorizont (94.4%) zeigen sich keine gegenwartsorientierte Verlaufsanpassungen. Das bedeutet, dass die weit überwiegende Mehrzahl der Prognosezeitreihen stärker die Zukunft als die Gegenwart reflektiert. Lediglich 10 von 72 Prognosezeitreihen (13.9%) weisen einen GOVA-Koeffizienten von > 1 auf. Diese Zeitreihen korrespondieren stärker mit den tatsächlichen Ereignissen zu den jeweiligen Zeitpunkten der Prognoseabgabe als mit den tatsächlichen Ereignissen zu den jeweiligen Zeitpunkten der Prognosegeltung. Sie sind also von gegenwartsorientierten Verlaufsanpassungen geprägt.

Prognosen sollten auf die Zukunft gerichtet sein. Prognosen, die eher die Gegenwart als die Zukunft widerspiegeln verfehlen im Prinzip ihren Zweck. Insofern sollte man erwarten, dass gegenwartsorientierte Verlaufsanpassungen im ökonomischen Prognosewesen krasse Ausnahmen darstellen. Jedoch zeigt sich, dass bei Finanzmarktprognosen gegenwartsorientierte Verlaufsanpassungen die Regel und nicht die Ausnahme sind. Eine ganze Reihe von Studien

⁶ Vgl. bspw. Spiwoks (2009).

gibt zu erkennen, dass mehr als 98% aller Finanzmarktprognosezeitreihen von gegenwartsorientierten Verlaufsanpassungen geprägt sind (Bofinger und Schmidt, 2003, Spiwoks, 2004, Spiwoks, Bedke und Hein, 2008, Spiwoks, Bedke und Hein, 2009, Spiwoks, Bedke und Hein, 2010, Spiwoks, Gubaydullina und Hein, 2011). Insofern muss es als Erfolg der Konjunkturprognostiker gewertet werden, dass nur ein kleiner Teil der Prognosezeitreihen von gegenwartsorientierten Verlaufsanpassungen geprägt ist.

Der Test auf Unverzerrtheit führt zu einer vollkommen anderen Bewertung des Prognoseerfolgs (Tabelle 4). Er vermittelt den Eindruck, dass die Anstrengungen der Prognostiker vollkommen gescheitert sind. Bei 31 der 36 betrachteten Prognosegegenstände (86.1%) zeigt der F-Test bei einer Irrtumswahrscheinlichkeit von 1%, dass $\alpha \neq 0$ und / oder $\beta \neq 1$ ist. Bei einer Irrtumswahrscheinlichkeit von 10% trifft dies sogar auf 33 von 36 Prognosegegenständen (91.7%) zu.⁷

Noch ungünstiger stellt sich die Situation dar, wenn man mit Hilfe des Durbin-Watson-Tests die Residuale auf systematische Komponenten hin untersucht. In allen 36 Fällen liegen bei einer Irrtumswahrscheinlichkeit von deutlich unter 1% Autokorrelationen vor.

Der Test auf Unverzerrtheit kennzeichnet somit die Prognosen zu allen 36 Prognosegegenständen als verzerrt.⁸ Die Prognoseanstrengungen sind, sofern man diesen Beurteilungsmaßstab als adäquat betrachtet, vollständig gescheitert. Keane und Runkle (1990) sowie Bonham und Cohen (2001) weisen darauf hin, dass Konsensprognosen als unverzerrt erscheinen können, obwohl die darin enthaltenen individuellen Prognosen durchaus verzerrt sind. Selbst dieser Effekt hat bei der vorliegenden Datenanalyse zu keinem günstigeren Ergebnis geführt.

Viele der betrachteten Prognosegegenstände genügen – abgesehen vom Test auf Unverzerrtheit – allen acht Beurteilungsmaßstäben (Tabellen 2 und 3):

⁷ Der Breusch-Pagan-Test (Breusch and Pagan, 1979) zeigt, dass bei 16 der 36 Datenbestände Heteroskedastizität angenommen werden muss. Bei diesen 16 haben wir den F-Test mit robusten Standardfehlern durchgeführt. Dies führte allerdings – bei Berücksichtigung von drei Nachkommastellen – nur in vier Fällen (fett hervorgehoben) zu einer Änderung des p-Wertes.

⁸ Ager, Kappler and Osterloh (2007) untersuchen unter anderem BIP-Prognosen für die meisten der hier betrachteten Länder im Zeitraum von 1996 bis 2006. Allerdings stellen sie in der Mehrzahl der Fälle fest, dass die Prognosen unverzerrt sind (S. 13). Dovern und Weisser (2009) untersuchen unter anderem BIP-Prognosen für die G7-Länder im Zeitraum von 1991 bis 2005. Auch sie kommen zu dem Ergebnis, dass die meisten Prognosen – insbesondere in Phasen ohne gravierende Strukturbrüche – unverzerrt sind (S. 20). Diese abweichenden Ergebnisse sind möglicherweise auf die kürzeren Untersuchungszeiträume, auf die abweichende Anzahl von Prognosen oder auf andere Zuschnitte des Tests auf Unverzerrtheit zurückzuführen. Batchelor (2007) untersucht unter anderem BIP-Prognosen für den Zeitraum von 1990 bis 2005. Für Japan, Italien, Deutschland und Frankreich erweisen sich die BIP-Prognosen als verzerrt (S. 20).

Die Prognosen zum BIP-Wachstum für die USA, Deutschland, Frankreich, Italien, Spanien, Kanada, die Niederlande und die Schweiz (a) erweisen sich für alle vier Prognosehorizonte (24, 21, 12 und 9 Monate) als den naiven Prognosen überlegen, (b) lassen sowohl anhand des absoluten Prognosefehlers wie auch des absoluten prozentualen Prognosefehlers bei einer Irrtumswahrscheinlichkeit von 1% einen Informationszuwachs im Zeitverlauf erkennen und (c) weisen bei keinem der beiden Prognosehorizonte (12 und 9 Monate) gegenwartsorientierte Verlaufsanpassungen auf.

Die Prognosen zum Wachstum der Industrieproduktion für die USA, Japan, Deutschland, Frankreich, Großbritannien, Italien und die Niederlande (a) erweisen sich für alle vier Prognosehorizonte (24, 21, 12 und 9 Monate) als den naiven Prognosen überlegen, (b) lassen sowohl anhand des absoluten Prognosefehlers wie auch des absoluten prozentualen Prognosefehlers bei einer Irrtumswahrscheinlichkeit von 1% einen Informationszuwachs im Zeitverlauf erkennen und (c) weisen bei keinem der beiden Prognosehorizonte (12 und 9 Monate) gegenwartsorientierte Verlaufsanpassungen auf.

Die Prognosen zum Wachstum des privaten Konsums für Deutschland, Großbritannien und Kanada (a) erweisen sich für alle vier Prognosehorizonte (24, 21, 12 und 9 Monate) als den naiven Prognosen überlegen, (b) lassen sowohl anhand des absoluten Prognosefehlers wie auch des absoluten prozentualen Prognosefehlers bei einer Irrtumswahrscheinlichkeit von 1% einen Informationszuwachs im Zeitverlauf erkennen und (c) weisen bei keinem der beiden Prognosehorizonte (12 und 9 Monate) gegenwartsorientierte Verlaufsanpassungen auf.

Bei 33 der 36 Prognosegegenstände (91.7%) werden folgende Anforderungen gleichzeitig erfüllt: (a) bei einem Prognosehorizont von neun Monaten werden die naiven Prognosen übertroffen, (b) bei Betrachtung des absoluten Prognosefehlers kann ein Informationszuwachs statistisch signifikant festgestellt werden und (c) bei einem Prognosehorizont von neun Monaten treten keine gegenwartsorientierte Verlaufsanpassungen auf (Tabelle 5). Lediglich die Prognosen zum Wachstum des privaten Konsums in der Schweiz und in Schweden sowie die Prognosen zum Wachstum der Industrieproduktion in Norwegen können diese drei Kriterien nicht vollständig erfüllen.

In Tabelle 5 zeigt sich auch die starke Diskrepanz zwischen den drei von uns bevorzugten Beurteilungsmaßstäben einerseits und dem Test auf Unverzerrtheit andererseits.

Wir stellen abschließend fest, dass Konjunkturprognostiker wesentlich erfolgreicher sind als Finanzmarktanalysten. Dieser gravierende Unterschied wird vom Test auf Unverzerrtheit allerdings nivelliert und verschleiert. Wir halten

den Test auf Unverzerrtheit deshalb für weitgehend ungeeignet, um ökonomische Prognosen zu beurteilen.

5 Fazit

Die vorliegende Studie setzt sich mit der Beurteilung von Konjunkturprognosen auseinander. Neben der Erörterung herkömmlicher Beurteilungsmaßstäbe werden auch zwei neue Verfahren vorgestellt: Der Test auf Informationszuwachs und der GOVA-Koeffizient.

Der Test auf Informationszuwachs betrachtet, ob die Prognosen bei allmählich rückläufigem Prognosehorizont genauer werden. Es wird also untersucht, ob die Prognostiker bereit und in der Lage sind, während des Prognosezeitraums relevante neue Informationen aufzunehmen und in ihren Prognosen zu berücksichtigen. Der GOVA-Koeffizient zeigt, ob die Prognosen stärker die Zukunft oder die Gegenwart reflektieren.

Wir untersuchen Konsensprognosen zur Entwicklung des BIP, der Industrieproduktion und der privaten Konsumausgaben in den USA, Japan, Deutschland, Frankreich, Großbritannien, Italien, Spanien, Kanada, Niederlande, Schweiz, Schweden und Norwegen.

Dabei stützen wir uns auf vier ausgewählte Beurteilungsverfahren: 1. Vergleich zur naiven Prognose mit Hilfe des mittleren absoluten Prognosefehlers bezogen auf die naive Prognose (*MRAPnP*), 2. Test auf Informationszuwachs, 3. GOVA-Koeffizient und 4. Test auf Unverzerrtheit.

Der mittlere absolute Prognosefehler bezogen auf die naive Prognose (*MRAPnP*) zeigt, dass die Konjunkturprognosen in 136 von 144 betrachteten Fällen (94,4%) exakter sind als die entsprechenden naiven Prognosen (Tabelle 2).

Auch der Test auf Informationszuwachs führt zu einer recht positiven Beurteilung der Konjunkturprognostiker. In 68 der 72 betrachteten Fälle (94,4%) zeigt sich eine statistisch signifikante Verringerung des Prognosefehlers bei allmählicher Verkürzung des Prognosehorizontes (Tabelle 3). In der weit überwiegenden Mehrzahl der Fälle sind die Konjunkturprognostiker also bereit und in der Lage, relevante Ereignisse während des sich allmählich verkürzenden Prognosehorizontes in ihren Prognosen zu berücksichtigen.

Die Analyse der Konjunkturprognosen mit Hilfe des GOVA-Koeffizienten führt ebenfalls zu einem günstigen Urteil. Bei 62 von 72 betrachteten Prognosezeitreihen (86,1%) liegen keine gegenwartsorientierten Verlaufsanpassungen vor (Tabelle 3). Die weit überwiegende Mehrzahl der Prognosen spiegelt also die Entwicklungen der Zukunft deutlich besser wider als die der Gegenwart.

Bei 33 der 36 Prognosegegenstände (91,7%) werden folgende Anforderungen gleichzeitig erfüllt: (a) bei einem Prognosehorizont von neun Monaten werden die naiven Prognosen übertroffen, (b) bei Betrachtung des absoluten

Prognosefehlers kann ein Informationszuwachs statistisch signifikant festgestellt werden und (c) bei einem Prognosehorizont von neun Monaten treten keine gegenwartsorientierte Verlaufsanpassungen auf (Tabelle 5). Dieses Ergebnis weist durchaus gut entwickelte Prognosefähigkeiten bei den Konjunkturexperten aus.

Der Test auf Unverzerrtheit betrachtet hingegen die untersuchten Prognosen ausnahmslos als gescheitert (Tabellen 4 und 5). Wir sind der Auffassung, dass die Diskrepanz zwischen dem Test auf Unverzerrtheit und den anderen drei Beurteilungsmaßstäben darauf zurückzuführen ist, dass der Test auf Unverzerrtheit unrealistische Ansprüche an die Prognostiker stellt.

Konjunkturprognosen sind wesentlich exakter als Finanzmarktprognosen. Dieser deutliche Vorsprung wird bei Anwendung des Tests auf Unverzerrtheit nivelliert und verschleiert. Wir plädieren deshalb dafür, dem Test auf Unverzerrtheit in Zukunft bedeutend weniger Aufmerksamkeit zu schenken als dies bisher der Fall ist.

Literatur

- Ackert, L. F. / Hunter, W. C. (1994), Rational Expectations and the Dynamic Adjustment of Security Analysts' Forecasts to New Information, in: *The Journal of Financial Research*, Bd. 17, S. 387-401.
- Ager, P / Kappler, M. / Osterloh, S. (2007), The Accuracy and Efficiency of the Consensus Forecasts: A Further Application and Extension of the Pooled Approach, Centre for European Economic Research Discussion Paper, Nr. 07-058, Mannheim.
- Andolfatto, D / Hendry S. / Moran, K. (2008), Are inflation expectations rational?, in: *Journal of Monetary Economics*, Bd. 55, S. 406-422.
- Andres, P. / Spiwoks, M. (1999), Forecast Quality Matrix, A Methodological Survey of Judging Forecast Quality of Capital Market Forecasts, in: *Journal of Economics and Statistics*, Bd. 219, S. 513-542.
- Ball, R. J. (1962), The Prediction of Wage-Rate Changes in the United Kingdom Economy 1957-60, in: *The Economic Journal*, Bd. 72, S. 27-44.
- Batchelor, R. (2007), Forecaster Behaviour and Bias in Macroeconomic Forecasts, ifo Working Papers, Nr. 39, ifo Institute of Economic Research at the University of Munich, Munich.
- Bofinger, P. / Schmidt, R. (2003), On the reliability of professional exchange rate forecasts: An empirical analysis for the EUR/US-\$ rate, in: *Financial Markets and Portfolio Management*, Bd. 17, S. 437-449.
- Bonham, C. S. / Cohen, R. H. (2001), To Aggregate, Pool, or Neither: Testing the Rational-Expectations Hypothesis Using Survey Data, in: *Journal of Business & Economic Statistics*, Bd. 19, S. 278-291.
- Bowles, C / Friz, R / Genre, V / Kenny, G / Meyler, A. / Rautanen, T. (2007), The ECB Survey of Professional Forecasters (SPF) – A Review After Eight Years' Experience, European Central Bank Occasional Paper Series, Nr. 59, April 2007, Frankfurt am Main.
- Breusch, T. S. / Pagan, A. R. (1979), Simple test for heteroscedasticity and random coefficient variation, in: *Econometrica*, Bd. 47, S.1287–1294.
- Brooks, R. / Gray, B. (2004), History of forecasters: Naive forecasts are better than the consensus WSJ forecast, in: *Journal of Portfolio Management*, Bd. 31, S. 113-117.
- Chinn, M. / Frankel, J. (1994), Patterns in Exchange Rate Forecasts for Twenty-five Currencies, in: *Journal of Money, Credit, and Banking*, Bd. 26, S. 759-770.
- De Gooijer, J. G. / Hyndman, R. J. (2006), 25 years of time series forecasting, in: *International Journal of Forecasting*, Bd. 22, S. 443-473.
- Diebold, F. X. (2007), Elements of Forecasting, 4th International Edition, Thomson Learning.

- Diebold, F. X. / Lopez, J. A. (1996), Forecast Evaluation and Combination, in: Maddala, G. S. / Rao, C. R. (Hrsg.), *Handbook of Statistics*. 14: S. 241-268, Amsterdam.
- Diebold, F. X. / Mariano, R. S. (1995), Comparing Predictive Accuracy, in: *Journal of Business and Economic Statistics*, Bd. 13, S. 253-263.
- Döpke, J. / Fritsche, U. (2006), Growth and inflation forecasts for Germany – A panel-based assessment of accuracy and efficiency, in: *Empirical Economics*, Bd. 31, S. 777-798.
- Dovern, J. / Fritsche, U. / Slacalek, J. (2009), Disagreement Among Forecasters in G7 Countries, European Central Bank Working Paper Series, Nr. 1082, Frankfurt am Main.
- Dovern, J. / Weisser, J. (2009), Accuracy, Unbiasedness and Efficiency of Professional Macroeconomic Forecasts: An Empirical Comparison for the G7, Jena Economic Research Papers, Nr. 2009-091, Friedrich Schiller University, Jena.
- Fair, R. C. / Shiller, R. J. (1990), Comparing Information in Forecasts from Econometric Models, in: *The American Economic Review*, Bd. 80, S. 375-389.
- Fildes, R. / Stekler, H. (2002), The state of macroeconomic forecasting, in: *Journal of Macroeconomics*, Bd. 24, S. 435-468.
- Fraser, P. / MacDonald, R. (1993), The Efficiency of CAC Stock Price Forecasts, in: *Revue économique*, Bd. 44, S. 991-1000.
- Harvey, D. I. / Leybourne, S. J. / Newbold, P. (1998), Tests for Forecast Encompassing, in: *Journal of Business & Economic Statistics*, Bd. 16, S. 254-259.
- Henriksson, R. D. / Merton, R. C. (1981), On Market Timing and Investment Performance, II. Statistical Procedures for Evaluating Forecasting Skills, in: *Journal of Business*, Bd. 54, S. 513-533.
- Joutz, F. / Stekler, H. O. (2000), An evaluation of the predictions of the Federal Reserve, in: *International Journal of Forecasting*, Bd. 16, S. 17-38.
- Keane, M. P. / Runkle, D. E. (1990), Testing the Rationality of Price Forecasts: New Evidence from Panel Data, in: *The American Economic Review*, Bd. 80, S. 714-735.
- Lakonishok, J. (1980), Stock Market Return Expectations: Some General Properties, in: *The Journal of Finance*, Bd. 35, S. 921-931.
- Laplace, P. S. (1814), *Essai philosophique sur les probabilités*, Paris.
- Mankiw, N. G. / Reis, R. / Wolfers, J. (2003), Disagreement About Inflation Expectations, in: *NBER Macroeconomics Annual*, Bd. 18, S. 209-248.
- Manzur, M. (1988), How Much Are Exchange Rate Forecasts Worth? in: *Australian Journal of Management*, Bd. 13, S. 93-113.

- Mathews, B. P. / Diamentopoulos, A. (1994), Towards a Taxonomy of Forecast Error Measures, A Factor-comparative Investigation of Forecast Error Dimensions, in: *Journal of Forecasting*, Bd. 13, S. 409-416.
- Mehra, Y. P. (2002), Survey Measures of Expected Inflation: Revisiting the Issues of Predictive Content and Rationality, in: *Federal Reserve Bank of Richmond Economic Quarterly*, Bd. 88/3, S. 17-36.
- Merton, R. C. (1981), On Market Timing and Investment Performance, I. An Equilibrium Theory of Value for Market Forecasts, in: *Journal of Business*, Bd. 54, S. 363-406.
- Mincer, J. / Zarnowitz, V. (1969), The Evaluation of Economic Forecasts, in: Mincer, J. (Ed.), *Economic Forecasts and Expectation*, New York, Columbia University Press, S. 3-46.
- Mose, J. S. (2005), Expert Forecasts of Bond Yields and Exchange Rates, in: *Danmarks Nationalbank Monetary Review*, Bd. 4th Quarter 2005, S. 91-95.
- Muth, J. F. (1961), Rational Expectations and the Theory of Price Movements, in: *Econometrica*, Bd. 29, S. 315-334.
- Osterloh, S. (2008), Accuracy and Properties of German Business Cycle Forecasts, in: *Applied Economics Quarterly*, Bd. 54, S. 27-57.
- Poulizac, D / Weale, M. / Young, G. (1996), The Performance of National Institute Economic Forecasts, in: *National Institute Economic Review*, Nr. 156, S. 55-62.
- Schuh, S. (2001), An Evaluation of Recent Macroeconomic Forecast Errors, in: *New England Economic Review*, Bd. Jan./Feb. 2001, S. 35-56.
- Simon, D. P. (1989), The Rationality of Federal F/s Rate Expectations: Evidence from a Survey, in: *Journal of Money, Credit, and Banking*, Bd. 21, S. 388-393.
- Scheier, J./ Spiwoks, M. (2006), Aktives Portfoliomanagement am britischen Anleihenmarkt, in: Meyer-Bullerdiel, F. / Spiwoks, M. (Hrsg.), *Reihe Bank- und Finanzwirtschaft*, Bd. 4.
- Spiwoks, M. (2004), The Usefulness of ZEW Stock Market Forecasts for Active Portfolio Management Strategies, in: *Journal of Economics and Statistics*, Bd. 224, 2004, S. 557-578.
- Spiwoks, M. (2009), Deceptive Unity: Surprising Characteristics of Pre-Tax Corporate Profit Forecasts, in: *International Research Journal of Finance and Economics*, Bd. 26, S. 170-177.
- Spiwoks, M / Bedke, N. / Hein, O. (2008), Forecasting the past: The case of US interest rate forecasts, in: *Financial Markets and Portfolio Management*, Bd. 22, S. 357-379.
- Spiwoks, M / Bedke, N. / Hein, O. (2009), The Pessimism of Swiss Bond Market Analysts and the Limits of the Sign Accuracy Test – An Empirical Investiga-

tion of Their Forecasting Success Between 1998 and 2007, in: *International Bulletin of Business Administration*, Bd. 4, S. 6-19.

Spiwoks, M / Bedke, N. / Hein, O. (2010), Topically Orientated Trend Adjustment and Autocorrelation of the Residuals - An Empirical Investigation of the Forecasting Behavior of Bond Market Analysts in Germany, in: *Journal of Money, Investment, and Banking*, Bd. 14, S. 16-35.

Spiwoks, M / Gubaydullina, Z. / Hein, O. (2011), Trapped in the Here and Now – New Insights into Financial Market Analyst Behavior, Wolfsburg Working Papers, Nr. 11-01, Ostfalia University of Applied Sciences, Wolfsburg.

Storbeck, O. (2010), Wie die Finanzkrise die VWL auf den Kopf stellt, in: *Handelsblatt*, 11. Januar 2010, S. 20-21.

Theil, H. (1971), *Principles of Econometrics*, New York.